Best
Available
Copy

AD-764 968

SPEECH AND PICTURE PROCESSING

Alan V. Oppenheim

Massachusetts Institute of Technology

Prepared for:

Office of Naval Research
Advanced Research Projects Agency

2 August 1973

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

Massachusetts Institute of Technology
Research Laboratory of Electronics
Cambridge, Massachusetts 02139

SPEECH AND PICTURE PROCESSING

Semiannual Technical Report
covering the period
January 15, 1973 - July 15, 1973

Sponsored by
Advance Research Project Agency
ARPA Order No. 1997

Contract No. N00014-67-A-0204-0064
Modification No: 01

Program Code: 80230 # 3D30

Effective Date of Contract:
January 15, 1973

Contract Expiration Date:
January 14, 1974

Amount of Contract:
$196, 377.00

Principal Investigator:
A. V. Oppenheim
617-253-4177

Scientific Officer:
Marvin Denicoff
Program Director for
Information Services

Semi-Annual Technical Report

ONR Contract N00014-67-A-0204-0064

covering the period January 16, 1973 - July 15, 1973

submitted by

A. V. Oppenheim - Principal Investigator

August 2, 1973

Projects Studied Under the Contract:

During the first half of the contract year, the program continued the following studies: speech analysis by linear prediction, digital frequency warping, development of a high speed digital processor for speech synthesis, and the design of two-dimensional recursive digital filters. These projects are summarized in the following pages and reprints of available publications are appended.

The views and conclusions contained in this document are those of the author, and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

1

Abstract

During the first half of the contract year, the program continued the following studies: speech analysis by linear prediction, digital frequency warping, development of a high speed digital processor for speech synthesis, and the design of two-dimensional recursive digital filters. These projects are summarized in the following pages and reprints of available publications are appended.

1. Speech Analysis by Linear Prediction

The general direction of this research is to study in detail, and to summarize in quantitave terms, the acoustic properties of pre-stressed consonants of General American English.

We have chosen to study the acoustic properties of these consonants under a highly controlled environment, where syntactic and semantic influences are constrained to a maximal degree and where the phonetic environment can be varied in a systematic way. We feel very strongly that, in order to study the acoustic properties of phonetic units and find the relationships between these acoustic properties and their underlying invariant attributes, the construction of such a controlled data-base is essential.

We have restricted ourselves to study only pre-stressed consonants for several reasons. First, a stressed consonant-vowel sequence seems to be universal among all languages. Studying pre-stressed consonants might provide a common ground where similar studies of other languages could be attempted and compared. Secondly, stressed syllables in an utterance are probably articulated with greater care and effort, thereby resulting in a robust acoustic signal where the extraction of acoustic parameters is more reliable. It might even be hypothesized that the acoustic properties of consonants are least distorted by the environment when they appear in stressed consonant-vowel (C-V) syllables. Acoustic invariants of consonants are more likely to be observed in such environment. Therefore, this phonetic environment might provide, in some sense, the clearest indication of the ideal relationship between the underlying invariant attributes and their acoustic correlates.

It is hoped that studies of this type will lead to a better understanding of the nature of language. This study also has immediate application in the recognition of speech by machine. Since stressed syllables seem to provide more reliable acoustic information, it may be hypothesized that one should first exercise considerable effort in the phonetic recognition of stressed vowels and consonants. One would also expect algorithms proposed for continuous speech to perform quite reliably in such an environment.

The acoustic analysis system has been implemented on the FDP-Univac computer facility at Lincoln Lab. The system can generate linear prediction spectra as well as various acoustic parameters. Spectra and parameters are computed every 5 msec and are available for on-line display and examination. They can also be stored on magnetic tapes for later use.

Several considerations have been weighted in the construction of the data-base.

1) Our interest is in studying stressed, C-V syllables in a controlled phonetic environment.

2) We would like to study them as they would appear in continuous speech.

3) We would like to eliminate, as much as possible, the linguistic influences.

4) It was also decided that a final consonant should be added to the syllables, since certain English vowels do not appear in the final position of a word.

5) The data-base should be general enough so that later studies of vowels and post-stressed consonants can be carried out on the same speech material.

6) Finally, talker-dependent phenomena, as well as variations within

a given talker, should also be accounted for.

The format of the speech material was finally decided to be a nonsense word hə'C_1VC_2 embedded in a carrier sentence, "Say --- again." The size of the data-base presents a non-trivial problem in terms of data storage. A rough calculation indicates that if we are to include all C-V combinations, a mere 10 repetitions of each combination will result in some seventy-four thousand utterances or approximately thirty hours of speech material. We are in the process of recording these utterances, excluding only those stressed, C-V combinations that do not occur in English. However, we only intend to study a realistic but substantial subset of the utterances in detail.

It was decided that we will use the Univac-FDP system for signal processing and use the TX-2 computer for storage and on-line data analysis. An interactive display and data retrieval system is presently under development on the TX-2. The system, when completed, will be able to display acoustic information in various forms. Collection of statistics over a large corpus of data can also be done with relative ease.

Preliminary data indicates that phonetic units have certain acoustic cues that are environment-dependent and others that are reasonably invariant. It is well known that when speech sounds are connected to form an utterance, there is a considerable degree of overlap in the encoding process so that influence of one phonetic unit can be found in its adjacent units. We feel that there is probably as much coarticulation effect on the consonants as there is on vowels, i.e., acoustic properties of consonants are quite dependent on its vowel environment. With the availability of additional data, we'll be able to draw more specific conclusions on these observations.

2. Application of Digital Frequency Warping to Unequal Resolution and Vernier Spectrum Analysis

The application of a technique, which has been referred to as digital frequency warping, to unequal bandwidth spectral analysis has been investigated. With this technique, a sequence is transformed into a second sequence in such a way that the Fourier transforms of the original and transformed sequences are related by a non-linear transformation of their frequency axes. An equal bandwidth analysis, carried out on the transformed sequence, then corresponds to an unequal bandwidth analysis of the original sequence. In many spectral analysis applications, it is desirable to have the analysis bandwidth change with frequency. For example, in the analysis of noise generated by mechanical systems for detecting potential failures it is often important to utilize proportional bandwidth or constant Q analysis so that the form of the spectrum is invariant under a time scaling of the signal, which might result, for example, from a change in the speed of the system. In other instances, such as sonar data analysis, it is desirable to analyze wide bandwidth data while obtaining high resolution at the low frequencies. In this case, it is generally desired to have the analysis bandwidth increase with frequency, but the exact form of the bandwidth as a function of frequency is not crucial. While the class of frequency transformations, afforded by digital frequency warping, does not permit proportional bandwidth analysis, it does permit an analysis in which the analysis bandwidth increases or decreases with frequency and also permits a vernier analysis. A method for comparing the bandwidth of the frequency function afforded by digital frequency warping, with a constant Q analysis, has been developed, and a detailed study of the comparison carried out.

One of the potential advantages of digital frequency warping lies in its relatively simple hardware implementation. This implementation corresponds basically to a cascade chain of first-order digital filters and thus an important consideration in the use of this technique is its sensitivity to arithemetic round-off noise. As a part of the study of this technique, an analysis of the arithemetic round-off effects has been carried out. It is shown, in particular, that these effects are not severe.

This work has been presented in a paper entitled, "Unequal Resolution and Vernier Spectrum Analysis," by C. Braccini and A. Oppenheim at the Florence seminar on digital filtering, Sept. 1972, and also at the International Conference on Signal Processing, University of Erlagen, Germany.

3. Development of a High Speed Digital Processor

Work has continued on the detailed design of the Black Box and on component selection. Due to expansion of the data word size to 24 bits, the chip count is now about 700. Two of eight circuit boards have been received, and final design and layout of the multiplier, which uses these boards, is proceeding. We are writing a computer program to optimize the propagation delay of the multiplier with respect to chip count, since we estimate about 165 MC10181 ALU chips will be needed for the 24 X 24 bit multiplier. Despite its size, the multiplier should still operate in approximately 100 nanoseconds. Considerable effort has been expended on the I/O interface design. Since the Black Box data word length is 24 bits, its instruction word length is 50 bits, and we plan to interface the machine to either 18 or 16 bit host computers, we have provided a flexible I/O system which is easy to use for both fractional and integer data, as well as instructions.

4. Two-Dimensional Filter Design

Recent work on the problem of two-dimensional recursive filter design has been in the area of developing an algorithm which can be used to select rational functions in two variables which approximate a desired filter frequency response in an optimum sense. The algorithm which has been developed is a straightforward extension of a one-dimensional algorithm which approximates an ideal function by rational functions. This algorithm is the differential correction algorithm. This is an iterative procedure which is described in E. W. Cheny's Introduction to Approximation Theory (p. 171). It can be proven to yield the best approximation to a given function in the Tchebycsev (minimax) sense in a region. Implementing the algorithm on a digital computer, however, constrains us to work on a finite point set rather than over a connected region. The number of points used is constrained by the amount of available computer memory and computation time,and to date the largest net of points has had 544 members. This was observed empirically to be sufficient for some two-dimensional design problems, but not for all.

The differential correction algorithm has also proven to be a useful technique for designing optimum one-dimensional recursive filters, and is worthy of further research in this context, particularly in the area of weighted approximations where the prescribed error tolerance from the ideal design is allowed to vary as a function of frequency.

In addition, we have discovered that if certain assumptions concerning the data to be filtered are made, a two-dimensional filtering problem may be transformed into a one-dimensional filtering problem by applying the one-projection theorem of Mersereau [MIT, ScD. thesis, 1973]. In this case, standard one-dimensional filter design techniques may be employed to solve the filtering problem.

PUBLICATIONS

1. "Acoustic Properties of Pre-Stressed Consonants," presented at the SUR Acoustic Phonetic workshop at Santa Barbara, California, March 1-2, 1973.

2. "Acoustic Properties of Pre-Stressed Consonants: A Program of Research," presented at the ARPA Graduate Student Conference at Monterey, California, July 3-6, 1973.

3. "Acoustic Properties of Pre-Stressed Consonants," doctoral Thesis proposal to be submitted to the Department of Electrical Engineering, M.I.T.

4. C. Braccini, A. Oppenheim, "Applications of Digital Frequency Warping to Unequal Resolution and Vernier Spectrum Analysis," Florence seminar on digital filtering, Sept. 1972 (Attached herewith as Appendix I).

5. C. Braccini, A. Oppenheim, "Applications of Digital Frequency Warping to Unequal Resolution and Vernier Spectrum Analysis," Conference on Signal Processing, University of Erlangen, Erlangen, Germany, April 4-6, 1973 (Attached herewith as Appendix II).

6. A. V. Oppenheim, "Some Digital Signal Processing Activities at M.I.T., Conference on Signal Processing, University of Erlangen, Erlangen, Germany, April 4-6, 1973 (Attached herewith as Appendix III).

7. A. V. Oppenheim, "Homomorphic Deconvolution," Proc. Seminaire G.U.T.S., Centre d'Etudes Nucléaires de Grenoble, Grenoble, France, June 14-15, 1973.

APPENDIX I

To be presented at Florence Seminar on Digital Filtering
September 21 to 22, 1972
Florence, Italy

Applications of Digital Frequency Warping to unequal Resolution and Vernier Spectrum Analysis

C. Braccini* and A. Oppenheim†

Recently, a technique referred to as digital frequency warping was proposed[1]. The technique corresponds to processing a sequence $f(n)$ to obtain a new sequence $g(n)$ in such a way that the Fourier transforms of the sequences are related by a distortion in the frequency axis. Specifically with $G(e^{j\hat{\Omega}})$ and $F(e^{j\Omega})$ denoting the Fourier transforms of $g(n)$ and $f(n)$ respectively the frequency variables are related by

$$\hat{\Omega} = \Omega + 2 \arctan \left[\frac{a\sin\Omega}{1-a\cos\Omega}\right] \qquad (1)$$

The sequence $g(n)$ is obtained from $f(n)$ by means of a cascade of two first order filters and a chain of first order all-pass networks all with pole locations at $z = a$. As the parameter a is varied, then according to eq. (1) the amount of frequency warping is varied. Because this cascade chain of first-order networks is defined by the single parameter a and because of its modular nature it leads to a relatively straightforward hardware realization.

The frequency warping specified by eq. (1) is a monotonic function of Ω and maps the interval $-\pi<\Omega<+\pi$ into the interval

$-\pi \leq \Omega \leq +\pi$. With the coefficient a positive, $\hat{\Omega}(\Omega)$ has a slope greater than unity at $\Omega=0$ and less than unity at $\Omega=\pi$. Consequently if constant bandwidth spectral analysis is carried out on the sequence g(n), using, for example, the Fast Fourier Transform algorithm, it will correspond to unequal bandwidth analysis of the sequence f(n) with the bandwidth monotonically increasing with frequency. If the relationship between Ω and $\hat{\Omega}$ were logarithmic then an equal bandwidth analysis of g(n) would result in a constant Q analysis of f(n). In many applications unequal resolution analysis is desirable[2-8]. In some cases the specific character of the resolution as a function of frequency is not important. In other applications such as when form invariance is desirable, it is important to have an approximately constant Q analysis. One of the objectives of this talk is to discuss the use of digital frequency warping to approximate constant Q analysis.

Because of the fact that the warping specified by eq. (1) has a slope greater than unity at $\Omega=0$ it can be used to provide a vernier spectrum analysis. In particular the Discrete Fourier Transform of g(n) provides a sampling of the spectrum of f(n) with close spacing of the spectral samples around $\Omega=0$. With the parameter a close to unity, this can be viewed as a vernier analysis around $\Omega=0$. In general, of course, we would like the ability to carry out a vernier analysis around an arbitrary frequency. This can be accomplished by multiplying f(n) by a complex exponential sequence to modulate the desired center frequency to $\Lambda=0$, implementing the frequency warping and Fourier transform. Several examples illustrating the procedure will be presented.

References

1. A. Oppenheim and D. Johnson "Discrete Representation of Signals" Proceedings of the IEEE vol. 60 No. 6, June 1972, pp 681-691.

2. J.S. Bendat, A.G. Piersol: "Design Considerations and use of analog power spectral density analyzers" Honeywell, Denver, Colo., Tech Rep., 1964.

3. C.D. Negron "Digital one-third octave spectral analysis" J.A.C.M. vol. 13, pp 605-614, 1866.

4. G.C. Maling, Jr., W.T. Morrey, W.W. Lang "Digital Determination of Third-Octave and Full-Octave Spectra of Acoustical Noise" IEEE Trans. on Au & Electr., vol. AU-15, No. 2, June 1967.

5. C.R. Arnold "Spectral Estimation for Transient Waveforms" IEEE Trans. on Audio and Audio and Electroacoustics, vol. AU-18, No. 3, Sept. 1970.

6. G. Gambardella "A Contribution to the Theory of Short-Time Spectral Analysis with Nonuniform Bandwidth Filters. IEEE Trans vol. CT-18 no. 4, July 1971, pp 455-460.

7. C. Braccini, C. De Cena, G. Gambardella "A Digital Vocoder Based on Spectral Analysis with Non Uniform Resolution.

Some Preliminary Results" presented at the XIX International Conference on Communications, Genova, Italy, October 1971.

8. G.A. Nelson, L.L. Pfeifer, R.C. Wood "High-Speed Octave Band Digital Filtering" IEEE Trans. on Audio and Electroacoustics Vol. AU-20, No. 1, March 1972.

* Istituto di Elettrotecnica dell'Universita di Genova, Italy. Visiting scientist at MIT Research Laboratory of Electronics with support of a NATO fellowship.

Department of Electrical Engineering and Research Laboratory of Electronics, MIT, supported in part by the National Science Foundation under grant GK-31353 and in part by the Advanced Research Projects Agency monitored by ONR under contract N00014-67-A-0204-0064. Presently a Guggenheim Fellow at Ecole Nationale Superieure d'Electronique et de Radioelectricite, Grenoble, France.

APPLICATIONS OF DIGITAL FREQUENCY WARPING TO UNEQUAL BANDWIDTH AND VERNIER SPECTRUM ANALYSIS

C. BRACCINI * A.V. OPPENHEIM **

* Istituto di Elettrotecnica dell'Universita di Genova, Italy. This work was carried out in part while a visiting scientist at M.I.T. Research Laboratory of Electronics with support of a NATO fellowship.

** Department of Electrical Engineering and Research Laboratory of Electronics, M.I.T. This work was supported in part by the National Science Foundation under grant Gk-31353 and in part by the Advanced Research Projects Agency monitored by ONR under contract N 00014-67-A-0204-0064. Presently a Guggenheim Fellow at Ecole Nationale d'Electronique et de Radioélectricité - Grenoble, France.

To be presented at the Conference on
SIGNAL PROCESSING
University of ERLANGEN - Germany
April 4 - 6 1973

1. INTRODUCTION

Spectral analysis has traditionally played an important role in the more general area of signal processing. With the development of the Fast Fourier Transform (FFT) algorithm and the present trend in digital hardware it has become increasingly possible to carry out sophisticated spectral analysis digitally.

In its most straightforward application in digital spectral analysis, the use of the FFT leads to spectrum analysis corresponding to an analysis carried out with a filter bank with equal bandwidth filters uniformly spaced over the entire signal band. In many spectral analysis applications, however, it is desirable to have the analysis bandwidth change with frequency.

For example in the analysis of noise generated by mechanical systems for detecting potential failures, it is often important to utilize proportional bandwidth, or constant Q analysis, to obtain form invariance. In other instances, it is desirable to analyze wide bandwidth data while obtaining high resolution at the low frequencies. In this case, it is generally desired to have the analysis bandwidth increase with frequency but the exact form of the bandwidth as a function of frequency is not crucial.

Another example of what can be regarded as spectrum analysis with an analysis bandwidth which is frequency dependent is a vernier analysis. In this case, we are interested in an analysis over a small portion of the band, with the effective filter spacing much smaller than the effective filter widths so that the spectrum is oversampled in frequency. Such as analysis is often useful, for example, when one is interested in detecting and measuring the center frequency of a narrowband component with a simple peak-picking algorithm on the spectrum. This corresponds to very high spectral sampling in a part of the frequency band, and none or very low spectral sampling in the remainder of the band.

In this paper, the application of a technique referred to as digital frequency warping to unequal bandwidth spectral analysis will be discussed. This technique transforms a sequence in such a way that the Fourier transforms of the original and transformed sequences are related by a nonlinear transformation of the frequency axis. An equal resolution analysis carried out on the transformed sequence then corresponds to an unequal resolution of the original sequence.

In the following discussion, we will first present a frame-work for discussing the problem of spectral analysis and indicate how the Discrete Fourier Transform is commonly applied. We will then discuss in a general sense how unequal bandwidth spectral analysis can be implemented by first implementing a nonlinear distortion of the frequency axis, followed by an equal bandwidth analysis.

2. GENERAL SPECTRAL ANALYSIS

Let $f(n)$ denote a sequence of data and $F(\Omega)$ its Fourier transform. A set of measurements of the spectrum as viewed through a spectral window will be denoted by G_k with

$$G_k = \left| \int_{-\pi}^{\pi} F(\Omega)\, H(\Omega,\Omega_k)\, d\Omega \right| \quad ; k = 1, 2, \ldots, N. \qquad (1)$$

The fact that the spectral window $H(\Omega,\Omega_k)$ is a function of two variables indicates that the shape and in particular the width of the window can change with the center frequency Ω_k. When the spectral window $H(\Omega,\Omega_k)$ depends only on the difference $(\Omega - \Omega_k)$ so that eq. (1) becomes

$$G_k = \left| \int_{-\pi}^{\pi} F(\Omega)\, H(\Omega - \Omega_k)\, d\Omega \right| , \qquad (2)$$

an equal bandwidth analysis results. The spectral measurements G_k can be thought of as corresponding qualitatively to filter bank outputs where each filters has the same spectral shape with only the center frequency Ω_k changing along the filter bank. Generally, it is desirable to choose the lowpass prototype filter characteristic $H(\Omega)$ to approximate unity over a band of frequencies with a width denoted by B and zero outside the band and to choose the spacing of the center frequencies i.e. $(\Omega_{k+1} - \Omega_k)$ to equal the constant B independent of k. In this way, the band is covered by non-overlapping filters. In some cases, however, it is desirable to choose the spacing of the center frequencies to be much less than the filter widths. In this case, the filters are overlapping. Such an analysis is commonly referred to as a vernier analysis and as mentioned in the introduction, can be useful if, for example, the analysis is directed toward the detection and measurement of a narrowband component.

If the spectral analysis corresponds to proportional bandwidth analysis then equation (1) takes the form

$$G_k = \left| \int_{-\pi}^{\pi} F(\Omega)\, H\left(\frac{\Omega}{\Omega_k} \right) d\Omega \right| . \qquad (3)$$

In this case, the effective filter width is proportional to the center frequency. Here the effective filters become wider as the center frequency increases.

Generally when data is analyzed using the discrete Fourier transform, a finite length window $w(n)$ is applied to the data and the DFT of the product computed. The magnitude of the DFT values then correspond to spectral samples G_k as specified by eq. (2) with $\Omega_k = k\,\Omega_0$. The Fourier transform of the data window $w(n)$ corresponds to the spectral window $H(\Omega)$. Thus, a direct application of the DFT results in an equal resolution analysis. A commonly used procedure for approximating a constant Q analysis or more generally unequal resolution is to sum adjacent frequency bins with the number of bins summed increasing with frequency. In the next section, we discuss an alternative whereby we consider the general notion of distorting the frequency axis of the signal to be processed, followed by an equal resolution analysis of the result.

3. UNEQUAL BANDWIDTH SPECTRUM ANALYSIS BASED ON FREQUENCY TRANSFORMATION

The DFT computes spectral samples whose magnitude is of the form

$$G_k = \left| \int_{-\pi}^{\pi} F(\Omega)\; H(\Omega - k\,\Omega_0)\; d\Omega \right| \tag{4}$$

corresponding to an equal bandwidth analysis of f(n). Let us consider a sequence $\hat{f}(n)$ which is related to f(n) in such a way that

$$F(\Omega) = \hat{F}(\hat{\Omega}) \tag{5}$$

where $\hat{\Omega} = \theta(\Omega)$.

An equal resolution analysis of $\hat{F}(\hat{\Omega})$ according to equation (4) leads to the coefficients

$$\hat{G}_k = \left| \int_{-\pi}^{\pi} \hat{F}(\hat{\Omega})\; H(\hat{\Omega} - k\,\hat{\Omega}_0)\; d\hat{\Omega} \right| \tag{6}$$

or, since $\hat{\Omega} = \theta(\Omega)$,

$$\hat{G}_k = \left| \int_{-\pi}^{\pi} F(\Omega)\; H\Big(\theta(\Omega) - \Omega_k\Big)\left(\frac{d\theta(\Omega)}{d\Omega}\right)\; d\Omega \right| \tag{7}$$

with $\Omega_k = k\theta(\Omega_0)$.

We see now, that this no longer has the form of an equal resolution analysis but, except for the factor $d\theta(\Omega)/d\Omega$, still retains the general form of equation (1). Thus, we can interpret equation (7) as an unequal resolution analysis of f(n) modified by the frequency characteristic $d\theta/d\Omega$. This spectral weighting can be compensated for prior to the spectral analysis by means of an appropriate pre-emphasis, or can be taken into account in interpreting the spectral coefficients. If the lowpass prototype filter $H(\Omega)$ has a bandwidth B, and if we assume that $\theta(\Omega)$ is approximately linear over any interval of length B, then the bandwidth of $H\big(\theta(\Omega) - k\theta(\Omega_0)\big)$ is B divided by the slope of $\theta(\Omega)$ at $\theta(\Omega) = k\theta(\Omega_0)$. Thus, if $\theta(\Omega)$ has slope monotonically decreasing with Ω, then the bandwidth of the analysis will monotonically increase with Ω.

For the specific case in which we would like a constant percentage bandwidth analysis, $\theta(\Omega)$ is of the form

$$\theta(\Omega) = a_1 \ln a_2 \Omega \tag{8}$$

so that eq. (7) takes the form

$$\hat{G}_k = \left| \int_{-\pi}^{\pi} F(\Omega)\; H\left(a_1 \ln \frac{\Omega}{\Omega_k}\right) \frac{a_1}{\Omega}\; d\Omega \right| \tag{9.a}$$

where :

$$\Omega_k = a_2^{(k-1)}\; \Omega_0^k \quad . \tag{9.b}$$

For a vernier analysis, we would choose $\theta(\Omega)$ to have a large slope in the frequency range for which the vernier analysis is desired.

There is not available a simple procedure for modifying a sequence to obtain a logarithmic frequency transformation of the form of eq. (8). However procedure has been developed for obtaining a frequency transformation of a specific form. In the next section, we briefly review this procedure and then discuss its use for unequal resolution analysis. As a measure of the behavior of the resolution as a function of frequency, we compare it with constant percentage bandwidth.

4. FREQUENCY WARPING USING AN ALL-PASS TRANSFORMATION

A procedure for implementing a frequency transformation of the form

$$\theta(\Omega) = (\Omega - \Omega_a) + 2 \arctan\left[\frac{a \sin(\Omega - \Omega_a)}{1 - a \cos(\Omega - \Omega_a)}\right] \tag{10}$$

has been discussed [1]. The frequency warping achieved is restricted to the functional form of eq. (10) and can only be varied within the flexibility afforded by the parameters $\underline{a}$ and Ω_0. A plot of $\theta(\Omega)$ for $\Omega_a = 0$ and several values of $\underline{a}$ is depicted in figure 1.

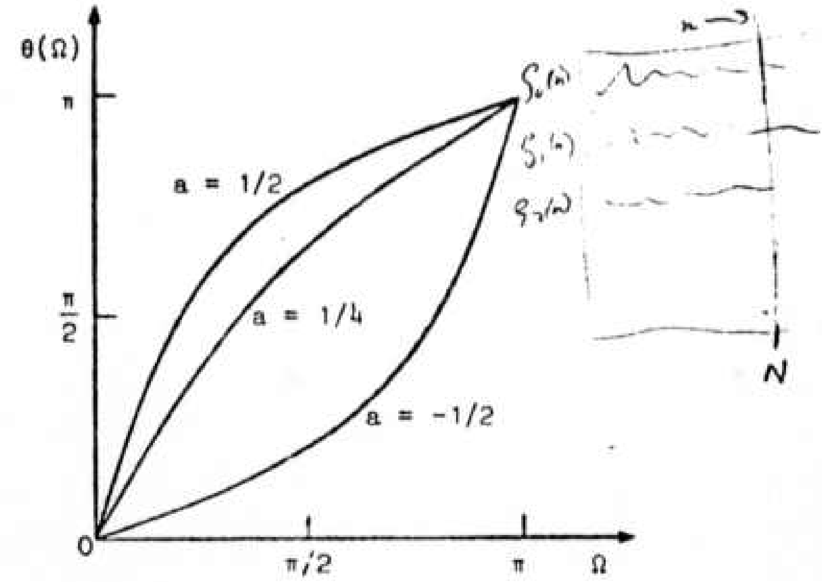


FIGURE 1.
Frequency warping function for several values of the parameter $\underline{a}$.

To implement this frequency warping, we consider two sequences $f(n)$ and $\hat{f}(n)$ with Fourier transforms related by $F(e^{j\Omega}) = \hat{F}(e^{j\theta(\Omega)})$.

With $\theta(\Omega)$ of the form of eq. (10), and assuming that $f(n)$ is zero for $n < 0$, it has been shown that $f(n)$ and $\hat{f}(n)$ can be related by

$$\hat{f}(k) = \sum_{n=0}^{\infty} e^{-j\Omega_a n} f(n)\; h_k(n) \tag{11}$$

where $H_k(z)$, the z-transform of $h_k(n)$ is given by

$$H_k(z) = \begin{cases} \dfrac{(1-a^2)z^{-1}}{(1-a\,z^{-1})^2}\left(\dfrac{z^{-1}-a}{1-a\,z^{-1}}\right)^{k-1} & k > 0 \\[2ex] \dfrac{1}{1-a\,z^{-1}} & k = 0 \end{cases} \tag{12}$$

this corresponds to processing $f(-n)e^{j\Omega_a n}$ with the system shown in figure 2.

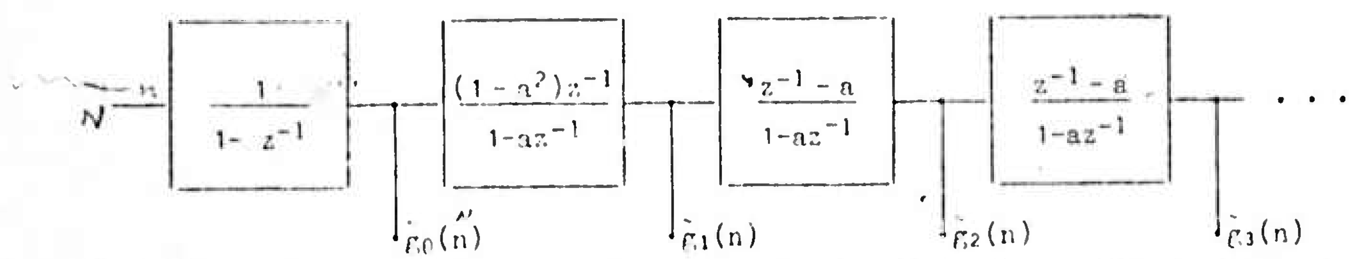


FIGURE 2. All pass network used to implement the frequency warping of fig. 1.

The sequence $\hat{f}(n)$ is obtained by sampling the outputs along the cascade chain at $n = 0$, i.e.

$$\hat{f}(n) = \tilde{g}_k(0) \qquad (13)$$

In a practical implementation of this system, the sequence $f(n)$ would be of finite length and it would often not be necessary to reverse the direction of the input sequence since this does not affect the magnitude of $F(e^{j\Omega})$ and $\hat{F}(e^{j\Omega})$ and only changes the sign of the phase of $F(e^{j\Omega})$.

For $\Omega_a = 0$ and a positive, we note from figure 1 that the slope is monotonically decreasing with frequency and consequently an analysis using the warping given by eq. (10) with $\Omega_a = 0$ will have an analysis bandwidth which increases with frequency. With $\Omega_a \neq 0$, the analysis bandwidth will be minimum around $\Omega = \Omega_a$. We wish now to compare the analysis as provided by the warping of eq. (10) with constant percentage bandwidth analysis. To present the basis for the comparison procedure it is helpful to refer to figure 3, indicating the manner in which a filter with bandwidth $\hat{B}$ is reflected by the curve $\theta(\Omega)$. The bandwidth of the filter applied to the sequence $\hat{f}(n)$ is

$$\hat{B} = \hat{\Omega}_2 - \hat{\Omega}_1 \qquad (14)$$

and its center frequency is

$$\hat{\Omega}_c = \frac{1}{2}\left(\hat{\Omega}_2 + \hat{\Omega}_1\right) \qquad (15)$$

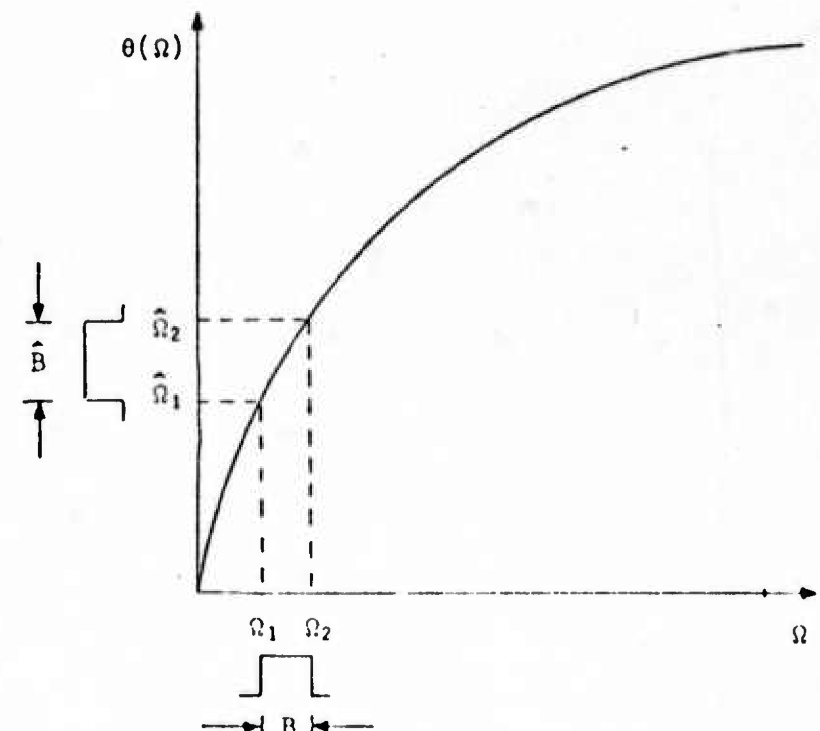


FIGURE 3.
The reflection of a filter with bandwidth $\hat{B}$ by the curve $\theta(\Omega)$.

The bandwidth and center frequency of the equivalent filter applied to $f(n)$ are

$$B = \Omega_2 - \Omega_1 \qquad (16)$$

$$\Omega_c = \frac{1}{2}\left(\Omega_2 + \Omega_1\right) \qquad (17)$$

The Q of the effective filters applied to the sequence $f(n)$ is defined as

$$Q = \frac{\Omega_c}{B} \qquad (18)$$

For a constant percentage bandwidth analysis we require that Q be independent of Ω_c and consequently that B be proportional to Ω_c. That this is in fact so for the logarithmic transformation $\hat{\Omega} = a_1 \ln a_2 \Omega$

or $\Omega = \frac{1}{a_2} \cdot e^{\hat{\Omega}/a_1}$

follows in a straightforward manner. Specifically, $\Omega_1 = \frac{1}{a_2} \exp\left(\frac{\hat{\Omega}_1}{a_1}\right)$, $\Omega_2 = \frac{1}{a_2} \exp\left(\frac{\hat{\Omega}_2}{a_1}\right)$ and $\hat{\Omega}_2 = \hat{\Omega}_1 + \hat{B}$. Therefore

$$Q = \frac{\Omega_c}{\Omega_2-\Omega_1} = \frac{1}{2} \frac{\exp\left(\frac{\hat{\Omega}_1}{a_1}\right)\left[\exp\left(\frac{\hat{B}}{a_1}\right)+1\right]}{\exp\left(\frac{\hat{\Omega}_1}{a_1}\right)\left[\exp\left(\frac{\hat{B}}{a_1}\right)-1\right]} = \frac{1}{2} \frac{\exp\left(\frac{\hat{B}}{a_1}\right)+1}{\exp\left(\frac{\hat{B}}{a_1}\right)-1} \quad . \tag{19}$$

Since $\hat{B}$ and $\underline{a}$ are constants, Q is a constant. In the same manner, we can compute the center frequencies and bandwidths reflected through the warping curve of eq. (10). The exact computation of the bandwidths leads to unwieldy expressions and we will instead base the comparison on an approximate analysis. The approximate analysis corresponds to assuming that $\theta(\Omega)$ is linear over an interval with width $\hat{B}$ in $\hat{\Omega}$. With this approximation, $\hat{B}$ and B can be simply related by the slope of $\theta(\Omega)$ at $\Omega = \Omega_c$ and $\hat{\Omega}_c$ and Ω_c are simply related by

$$\hat{\Omega}_c = \theta(\Omega_c) \quad . \tag{20}$$

It has been demonstrated that for a = 1/2 and $0 \le \hat{B} \le 2\pi/32$ the error due to this approximation is less than 1 % and for $\hat{B}$ less than $2\pi/16$ the error is less than 4 %. With the assumption that $\theta(\Omega)$ is linear over the range $\hat{\Omega}_2 - \hat{\Omega}_1$ with slope $\theta'(\Omega_c)$, $\hat{B}$ and B are simply related by

$$\hat{B} = B\,\theta'(\Omega_c) \quad . \tag{21}$$

Thus, the Q of the effective filters is

$$Q = \frac{\Omega_c}{B} = \frac{\Omega_c}{\hat{B}}\,\theta'(\Omega_c) \quad . \tag{22}$$

For constant percentage bandwidth, Q is constant, while the Q given in eq. (22) is a function of Ω_c. To compare this with constant Q, we define the error relative to a constant value Q_r as

$$\varepsilon(\Omega_c) = \frac{Q(\Omega_c) - Q_r}{Q_r} \tag{23}$$

where we have denoted the dependance of the error and Q on Ω_c. With the error expressed in this way, Q_r is of course arbitrary and the error is therefore dependent on what constant Q value we choose to compare $Q(\Omega_c)$ with. Consequently, it is convenient to display the error in terms of $\log(1+\varepsilon(\Omega_c))$ since

$$\log(1+\varepsilon(\Omega_c)) = \log Q(\Omega_c) - \log Q_r \quad . \tag{24}$$

With $\varepsilon(\Omega_c)$ presented in this way a change in Q_r is reflected by a vertical displacement of the error curve. Since, from eq. (22), $Q(\Omega_c)$ is proportional to $1/\hat{B}$, it is also convenient to consider Q_r to be proportional to $1/\hat{B}$. Thus we express Q_r as $Q_r = C/\hat{B}$.

In figure 4, a set of error curves are displayed for different values of $\underline{a}$ and Ω_a. For each value of $\underline{a}$, the different numbered curves presented correspond to a change in Ω_a. It was also convenient for the purpose of displaying these curves, to adjust the vertical position of each curve so that they are tangent to $\varepsilon = 0$. Thus each curve displayed has associated with it a different value of C.

A table of the values of Ω_a and C for each curve is given in Table 1. We note on these curves that for each value of a, those curves which are simple convex curves appear to be quite similar. It has been verified computationally that for values of modulation between those given, linear interpolation provides a close approximation. To discuss how the error curves should be interpreted, let us consider a specific example. Let us assume that we wish to investigate the deviation from constant Q for a = 1/2. Referring to the appropriate set of error curves, we note that from curve 3 (figure 4.b), corresponding to a modulation of zero, it is possible to have a positive error relative to $Q_r = 0.90\ \hat{B}$ between 0 and 10 % in the frequency range $0.42\ \pi \le \Omega \le 0.93\ \pi$. Alternatively by shifting the same curve downward, we see that it is possible to have an error between ± 10 % relative to $Q_r = \hat{B}$ in the frequency range $0.26\ \pi \le \Omega \le 1.04\ \pi$. For a = 3/4; with $\Omega_a = 0$, it is possible to have an error between ± 10 % in the frequeny range $0.43\ \pi \le \Omega \le 1.05\ \pi$ with $Q_r = 0.43\ \hat{B}$. We note that for Q_r to be same for the two different values of a , $\hat{B}$ can be larger if a = 1/2 than if a = 3/4. Thus, with a = 3/4; a wider bandwidth analysis can be used on the sequence $\hat{f}(n)$.

In a practical setting, we may wish to use the frequency warping to implement an approximation to a constant Q analysis. Given the frequency range to be analyzed and the allowable deviation from constant Q, we can then use the curves of figure 4 to choose the parameters a and Ω_a. The design procedure will generally be of a trial and error kind. The choice of Ω_a will generally depend on the range to be analyzed. The parameter a influences tha value of Q_r so that as a increases the value of $\hat{B}$ required to approximate a given Q will decrease requiring a finer resolution analysis of $\hat{f}(n)$. At the same time, it is generally possible to extend the frequency range to lower frequencies as a increases.

TABLE 1.

	a	curve No.	Ω_a	C
Figure 4.a	1/4	1	0.25	1.4
		2	0.55	1.9
		3	1	2.2
		4	1.5	2.5
		5	2	2.9
Figure 4.b	1/2	1	- 0.5	0.66
		2	- 0.32	0.75
		3	0	0.90
		4	0.5	1.1
		5	1.0	1.2
		6	1.5	1.4
		7	2.0	1.6
Figure 4.c	3/4	1	- 0.53	0.29
		2	- 0.25	0.35
		3	0	0.39
		4	0.5	0.47
		5	1	0.55
		6	1.5	0.63
		7	2	0.70
Figure 4.d	15/16	1	- 0.57	0.064
		2	- 0.4	0.073
		3	0	0.089
		4	0.5	0.11
		5	1	0.12
		6	1.5	0.14
		7	2	0.16

FIGURE 4.

Error curves for the use of digital frequency warping to approximate constant percentage bandwidth

FIGURE 4.a.
a = 1/4

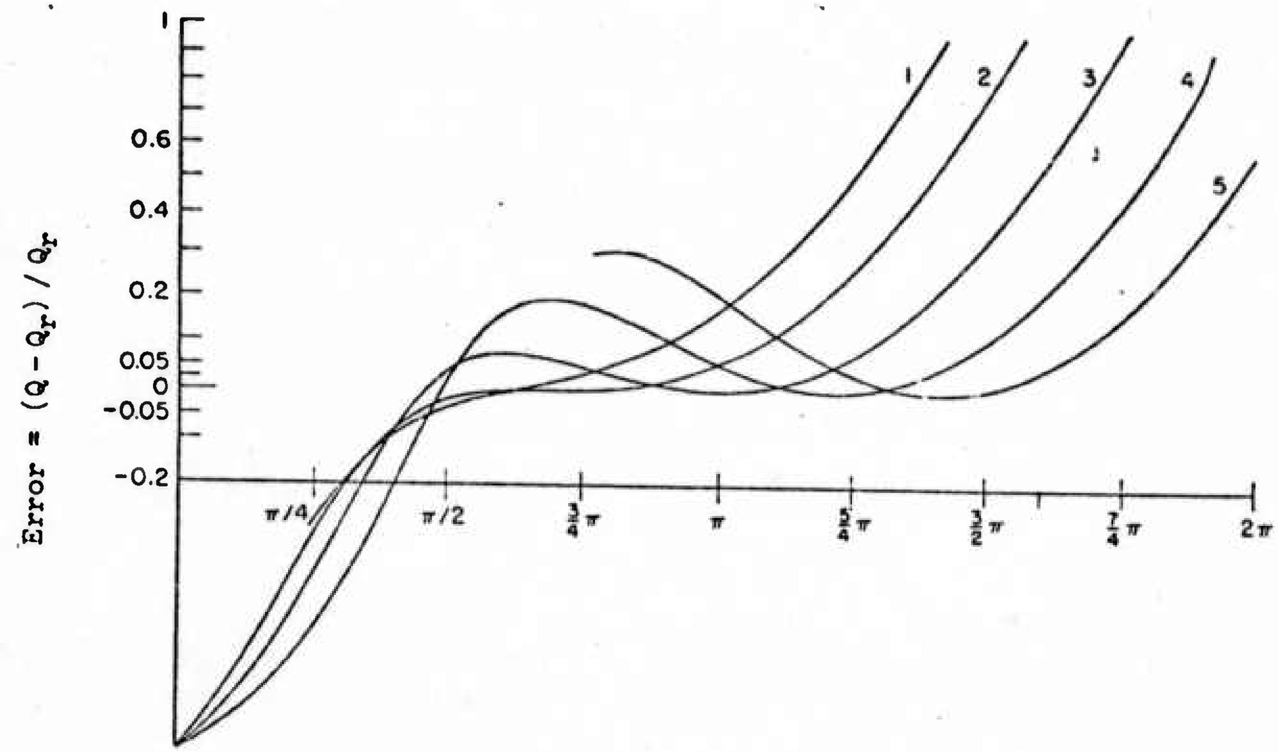


FIGURE 4.b.
a = 1/2

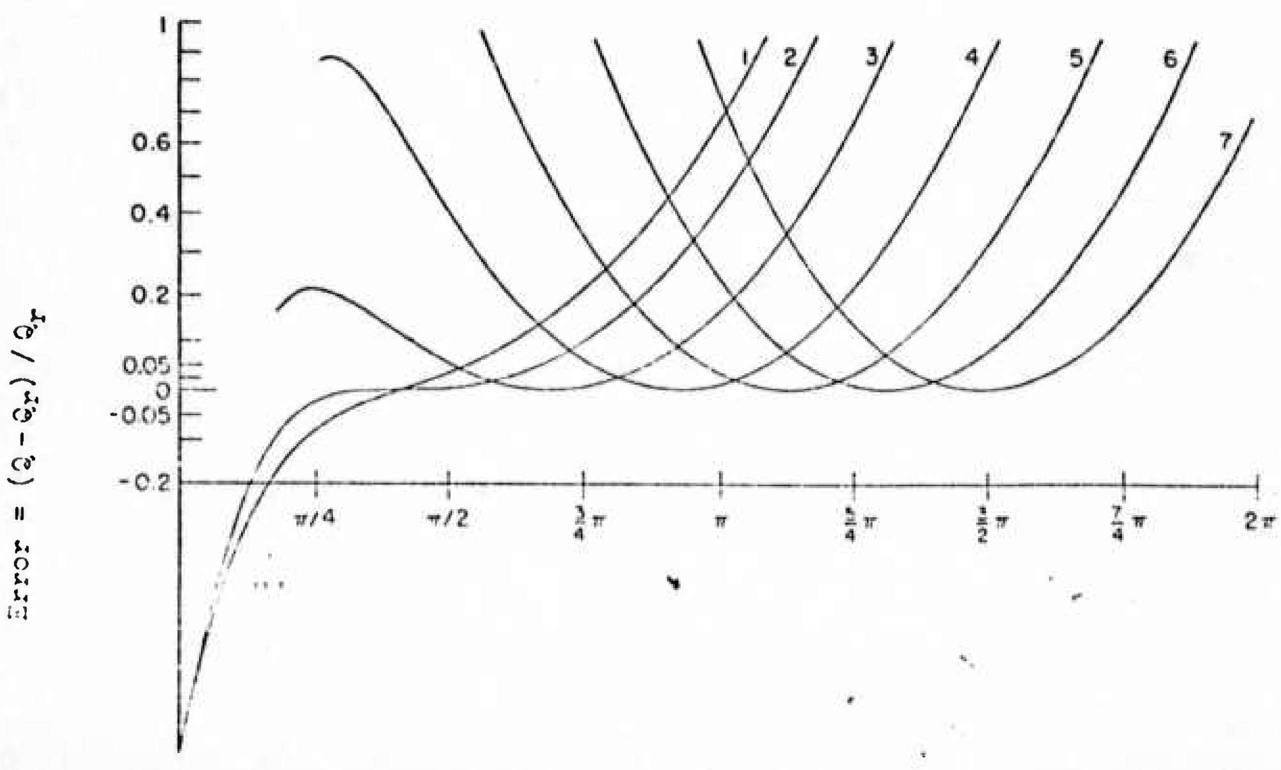


FIGURE 4.c.
a = 3/4

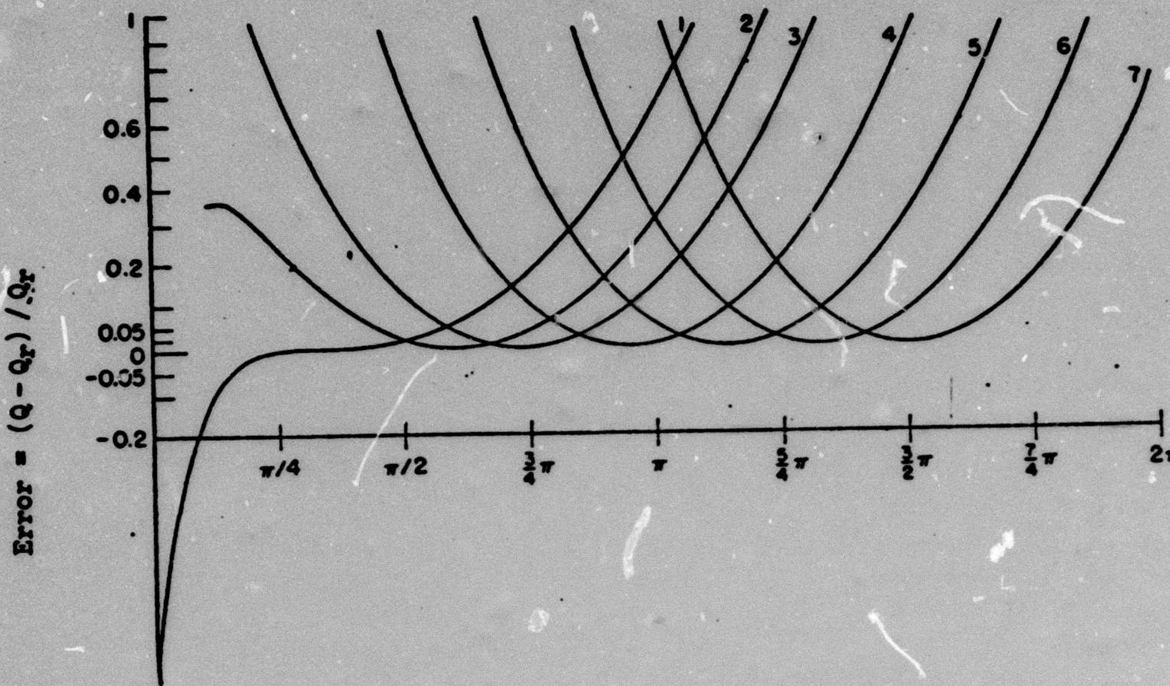
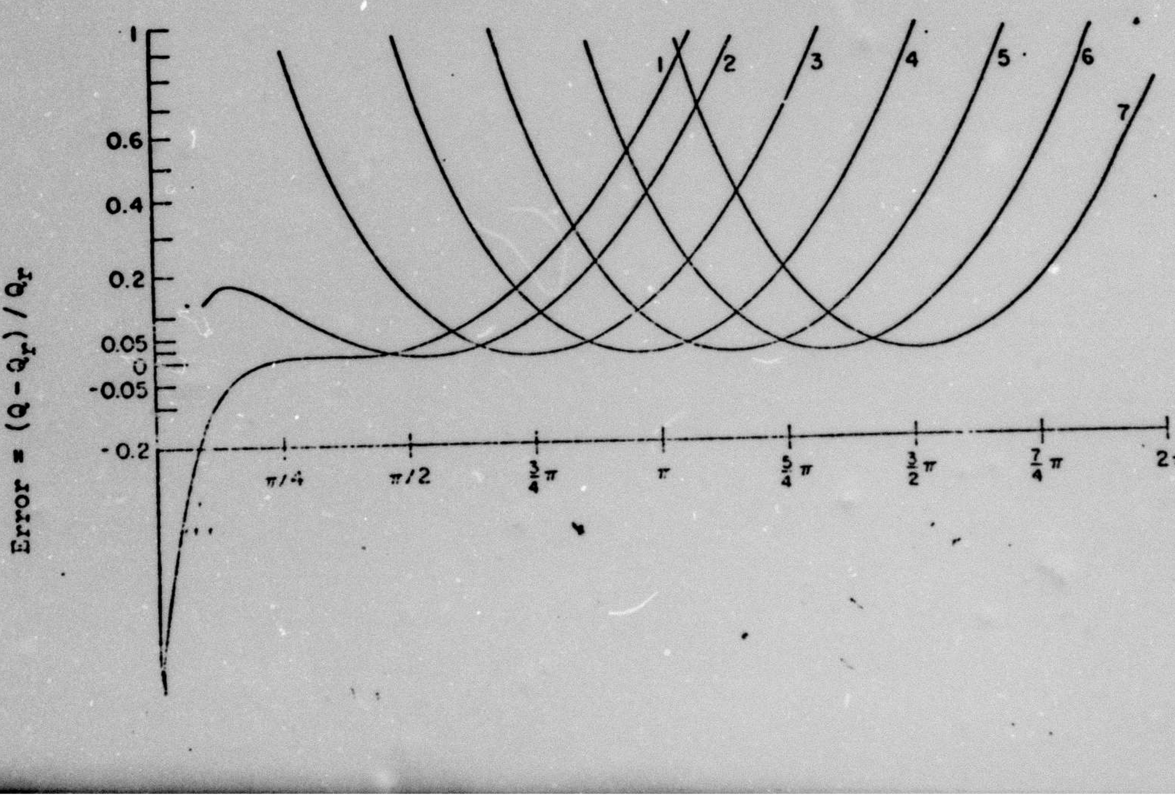


FIGURE 4.d.
a = 15/16



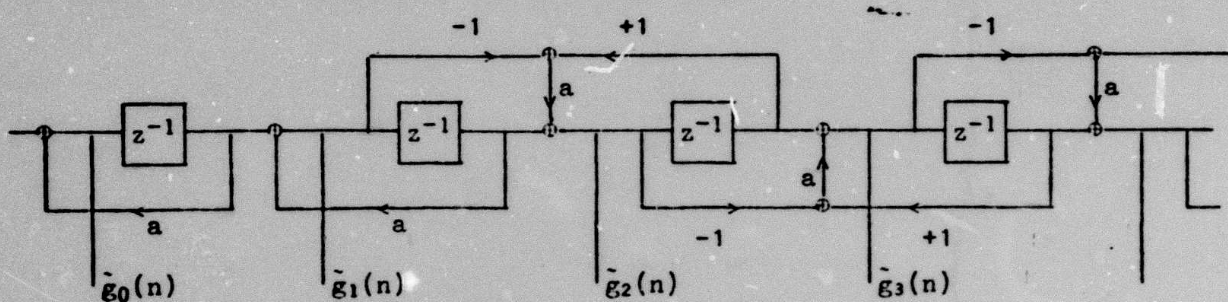
5. NOISE ANALYSIS OF DIGITAL FREQUENCY WARPING

The implementation of digital frequency warping utilizes a cascade chain of all-pass networks as depicted in figure 2.

In the previous sections, we have considered the effect of the parameter a with regard to the applications to unequal bandwidth analysis. From figure 2, we see that as the parameter a increases toward unity, the poles approach the unit circle. Consequently we would expect that with finite register length arithmetic considerations of roundoff noise will play an important role. Thus, this section is directed toward an analysis of roundoff noise in implementing digital frequency warping.

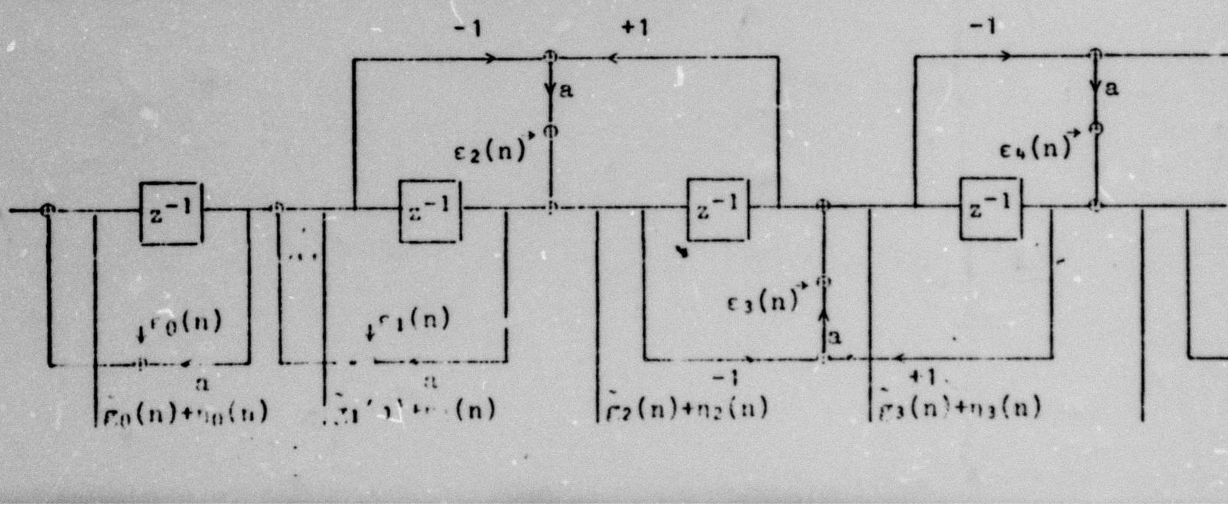
An implementation of the system of figure 2 in terms of multipliers, delays and adders is shown in figure 5. Note that, after the first delay, the outputs of this system differ from those of figure 2 by a factor of $(1-a^2)$. Since the system of figure 5 implies a simple hardware structure, it is the system of fig. 5 that will be analyzed with regard to roundoff noise. The factor $(1-a^2)$ is easily accounted for in a number of ways such as dividing $\hat{f}(0)$ by $(1-a^2)$ and applying a scale factor of $(1-a^2)$ in interpreting the amplitude of the sequence $\hat{f}(n)$.

FIGURE 5.
Representation of warping network
in terms of multipliers, delays and adders.



In analyzing the roundoff noise, we will consider only fixed point arithmetic and assume that rounding is applied after multiplication. The register length will be considered to be b-bits plus sign so that the error due to rounding is between plus and minus $\frac{1}{2}\cdot 2^{-b}$. We will assume that the rounding error can be represented statiscally. Thus, white noise generators can be inserted after the multiplies to account for roundoff noise as shown in figure 6.

FIGURE 6.
Network of figure 5 with roundoff noise sources inserted.



Each of the noise generators $\varepsilon_k(n)$ are assumed to be uncorrelated, white, uniformly distributed in amplitude between $\pm \frac{1}{2}\cdot 2^{-b}$ with variance : $\sigma_\varepsilon^2 = \frac{1}{12}\cdot 2^{-2b}$. The total noise output at the k^{th} tap is denoted by $\eta_k(n)$.

Let $H_{k,r}(z)$ denote the transfer function from the k^{th} noise source input to the r^{th} output tap where $r \geq k$. Then, since we assume that the $\varepsilon_k(n)$ are white and uncorrelated

$$\sigma_\eta^2(r) = \mathrm{var}\big(\eta_r(n)\big) = \sigma_\varepsilon^2 \sum_{k=0}^{r} \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{k,r}(e^{j\omega})|^2 \, d\omega \quad . \tag{25}$$

It can be verified by inspection of figure 6 that

$$\begin{aligned} H_{0,0}(z) &= \frac{1}{1 - a\, z^{-1}} \\ H_{0,r}(z) &= \frac{z^{-1}}{(1 - a\, z^{-1})^2} \left(\frac{z^{-1} - a}{1 - a\, z^{-1}} \right)^{r-1} & r \geq 1, \\ H_{k,r}(z) &= \frac{1}{1 - a\, z^{-1}} \left(\frac{z^{-1} - a}{1 - a\, z^{-1}} \right)^{r-k} & k \geq 1,\ r \geq k. \end{aligned} \tag{26}$$

Thus $$\frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{0,0}(e^{j\omega})|^2 \, d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{k,r}(e^{j\omega})|^2 \, d\omega = \frac{1}{1 - a^2} \qquad k \geq 1,\ r \geq k,$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{0,r}(e^{j\omega})|^2 \, d\omega = \frac{a^2(1 + a^2)}{(1 - a^2)^3} \qquad r \geq 1 .$$

Thus
$$\sigma_\eta^2(0) = \frac{1}{1 - a^2}\,\sigma_\varepsilon^2 \tag{27.a}$$

$$\sigma_\eta^2(r) = \left[\frac{a^2(1 + a^2)}{(1 - a^2)^3} + r\,\frac{1}{1 - a^2} \right] \sigma_\varepsilon^2 \quad . \tag{27.b}$$

It is interesting to note that the increase of $\sigma_\eta^2(r)$ with r is relatively mild. If we assume that a is not close to unity then for r large $\sigma_\eta^2(r)$ is approximately proportional to r. In addition to computing the output noise variance, however, we must consider the effect of limited dynamic range. With limited dynamic range taken into account, we can compute a noise to signal ratio at each output node in the network of figure 6.

To develop an approximate analysis, let us consider an input signal which is gaussian white noise with variance σ_f^2. Then the signal at each node is also gaussian. With $\sigma_s^2(r)$ denoting the variance of the output signal at the r^{th} node, it follows that

$$\sigma_s^2(r) = \sigma_f^2\,\frac{1}{2\pi}\int_{-\pi}^{\pi} |H_{0,r}(e^{j\omega})|^2 \, d\omega \quad . \tag{28}$$

Thus
$$\begin{aligned} \sigma_s^2(0) &= \sigma_f^2\,\frac{1}{1 - a^2} \\ \sigma_s^2(r) &= \sigma_f^2\,\frac{a^2(1 + a^2)}{(1 - a^2)^3} & r \geq 1 . \end{aligned} \tag{29}$$

For $a > 1/\sqrt{3}$, $\sigma_s^2(1) > \sigma_s^2(0)$. We note furthermore that $\sigma_s^2(r)$ is independent of r for r > 0. Thus let us assume that the maximum output signal variance occurs for r = 1. Since we have assumed that the signal is gaussian, the system will overflow a certain percentage of the time and as we scale the input signal the output signal variance is scaled and consequently also so is the overflow rate. It is generally convenient to consider the overflow rate to be independent of $\underline{a}$ in which case we scale σ_f such that $\sigma_s^2(r)$ is constant, i.e. $\sigma_s^2(r) = C$. In that case, the maximum output noise to signal ratio is given by

$$\frac{\sigma_n^2}{\sigma_s^2} = \frac{1}{C}\,\sigma_\epsilon^2 \left[\frac{a^2(1+a^2)}{(1-a^2)^3} + r\,\frac{1}{1-a^2} \right] \tag{30}$$

or, since $\sigma_\epsilon^2 = \frac{1}{12}\cdot 2^{-2b}$
$$\frac{\sigma_n^2}{\sigma_s^2} = \frac{1}{C}\,\frac{1}{12}\cdot 2^{-2b} \left[\frac{a^2(1+a^2)}{(1-a^2)^3} + r\,\frac{1}{1-a^2} \right] . \tag{31}$$

According to eq. (31), the noise to signal ratio increases monotonically with r and thus the largest noise-to-signal ratio occurs in the last stage. With R + 1 denoting the total number of stages, and assuming that R is relatively large, the noise-to-signal ratio in the last stage can be approximated as

$$\frac{\sigma_n^2}{\sigma_s^2} = \frac{1}{C}\,\frac{1}{12}\cdot 2^{-2b}\,\frac{R}{1-a^2} . \tag{32}$$

If, for example, we consider $a = 3/4$ and $R = 2^9$

$$\frac{\sigma_n^2}{\sigma_s^2} = \frac{1}{C}\,\frac{1}{12}\cdot 2^{-2b}\,\frac{2^{13}}{7} .$$

In general, we would choose the constant C on the basis of the percentage of overflow which we are willing to permit. For this example, let us choose C so that the standard deviation of the output is 1/4, i.e. C = 1/16. Then

$$\frac{\sigma_n^2}{\sigma_s^2} = \frac{2^{17}}{84}\cdot 2^{-2b} \quad ; \text{ or, in db, } \quad 10 \log_{10}\left(\frac{\sigma_n^2}{\sigma_s^2}\right) = 31.8 - 6\,b .$$

Thus for example with b = 17, the noise-to-signal ratio is approximately minus 70 db. If we choose C instead so that the standard deviation is 1/2, then with b = 17, the noise-to-signal ratio is minus 76 db.

The above analysis is based on the attitude that we are willing to permit the filter to overflow a certain percentage of the time. A more pessimistic analysis can also be carried out for which we scale the input to prevent overflow in the worst case. To develop this analysis, let $h_{0,r}(n)$ denote the impulse response of the system from the input to the k^{th} output node and $S_k(n)$ denote the output at the k^{th} node.

In general, the input will be of finite length N, and thus we assume that $x(n) = 0$ for $n < 0$ and $n \geq N$. Then
$$S_k(n) = \sum_{r=0}^{n} h_{0,k}(r)\;x(n-r)$$

and thus
$$|S_k(n)| \leq \max(x(n)) \sum_{r=0}^{n} |h_{0,k}(r)| . \tag{33}$$

We are interested in the system outputs for $n = N$ and thus we iterate the system N times. Then, for $n \leq N$, we write that

$$|S_k(n)| \leq \max_{\text{over } n}(x(n)) \sum_{r=0}^{N} |h_{0,k}(r)| \qquad n \leq N . \tag{34}$$

To prevent overflow, we require that $|S_k(n)| < 1$ for all k, which, from eq.(34), is guaranteed by scaling the input amplitude so that

$$\max_{\text{over } n} \left(x(n) \right) < \frac{1}{M}$$

where

$$M = \max_{\text{over } k} \sum_{r=0}^{N} |h_{0,k}(r)| \quad . \tag{35}$$

If we assume that the input is white noise uniformly distributed in amplitude between plus and minus $1/M$, then $\sigma_x^2 = 1/3M^2$ and the variance of the output signal $S_k(n)$ is

$$\sigma_s^2(0) = \frac{1}{3M^2} \frac{1}{1-a^2} \tag{36.a}$$

$$\sigma_s^2(k) = \frac{1}{3M^2} \frac{a^2(1+a^2)}{(1-a^2)^3} \qquad k \geq 1 \, . \tag{36.b}$$

Combining eqs. (36) and (27), the output noise-to-signal ratio is

$$\frac{\sigma_n^2(0)}{\sigma_s^2(0)} = 3 \; M^2 \; \sigma_\varepsilon^2 \tag{37.a}$$

$$\frac{\sigma_n^2(k)}{\sigma_s^2(k)} = 3 \; M^2 \left(1 + k \frac{(1-a^2)^2}{a^2(1+a^2)} \right) \sigma_\varepsilon^2 \qquad k \geq 1 \, . \tag{37.b}$$

From eqs. (37.a) and (37.b), it is clear that the noise-to-signal ratio is a monotonic function of k and consequently the largest noise-to-signal ratio occurs in the last stage. Then, with $K+1$ denoting the number of stages in the system, and assuming $K \gg 1$, the noise-to-signal ratio in the last stage can be approximated as

$$\frac{\sigma_n^2(K)}{\sigma_s^2(K)} = 3 \; M^2 \; K \; \frac{(1-a^2)^2}{a^2(1+a^2)} \; \sigma_\varepsilon^2 \quad . \tag{38}$$

To observe how the noise-to-signal ratio is influenced by the parameters K, a and N, we must consider the influence of these parameters on M as defined by eq. (35). To this end, we consider S(k) defined as

$$S(k) = \sum_{r=0}^{N} |h_{0,k}(r)| \tag{39}$$

so that

$$M = \max_{\text{over } k} \left(S(k) \right) \quad .$$

Figure 7 shows a plot of S(k) for several different values of a as a function of N and k. There are a number of trends evident from these plots. First, we note that for a given N and a, S(k) reaches a maximum and then decreases. Furthermore, this maximum is reached for k small compared with N, and the value of k for which S(k) reaches a maximum decreases as a increases. Consequently, it is reasonable to consider M to be independent of K in eq. (38). Thus, from eq. (38), the noise-to-signal ratio does not increase rapidly with K, i.e. it is relatively insensitive to the number of stages.

[1] A.V. OPPENHEIM and D.H. JOHNSON "Discrete Representation of Signals" Proceedings of the I.E.E.E., 60 n° 6, pp. 681-691 (june 1972).

FIGURE 7.
S(k) for several values of a as a function of N and k.

FIGURE 7.a.
a = 1/2

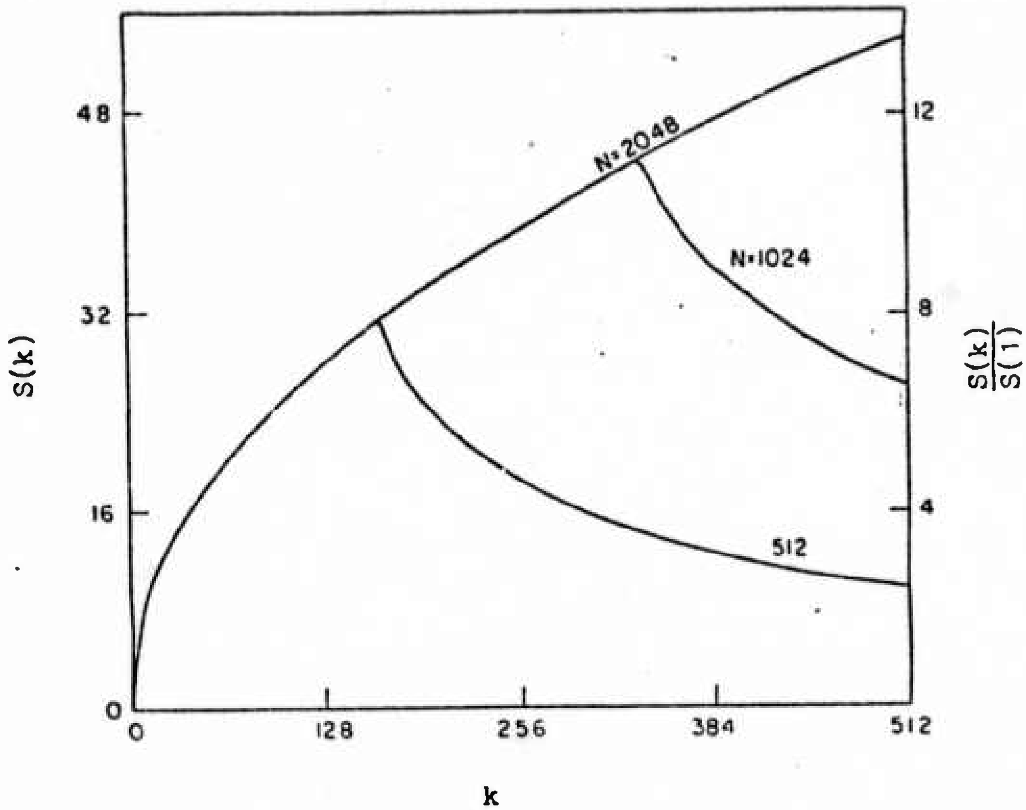


FIGURE 7.b.
a = 3/4

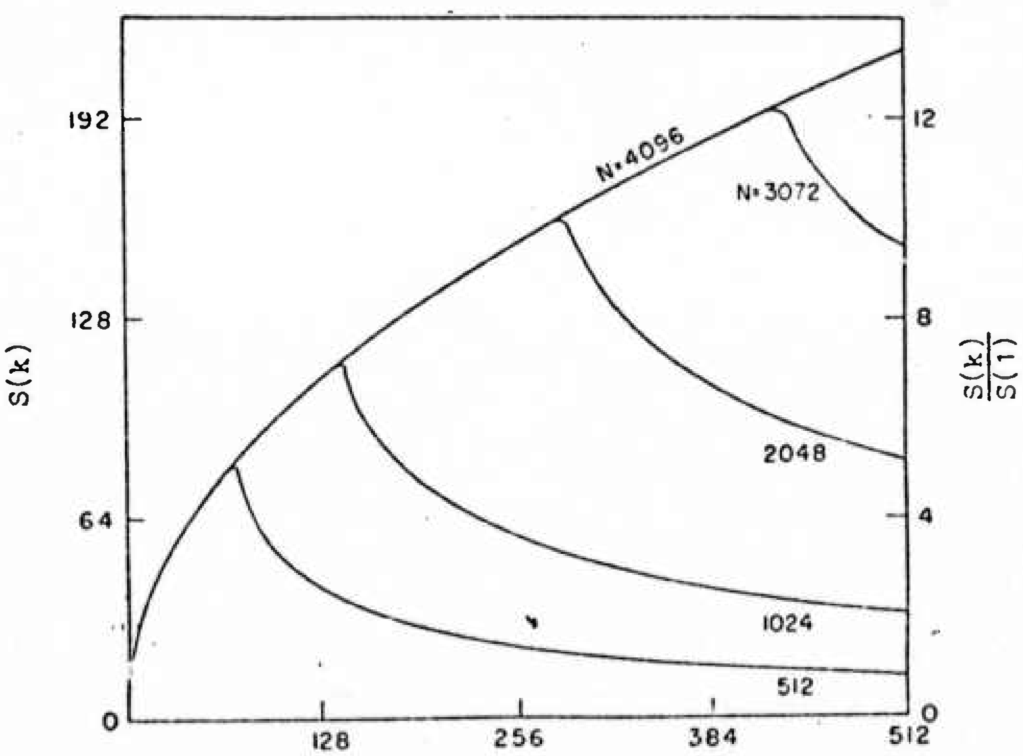
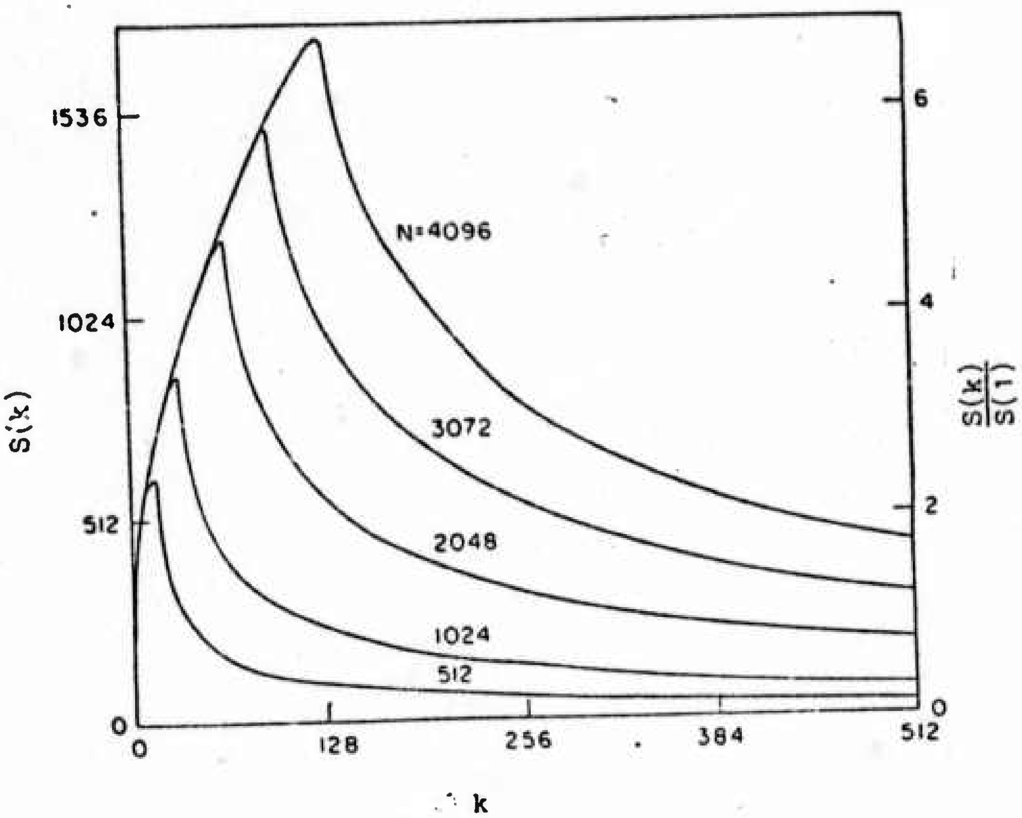


FIGURE 7.c.
a = 15/16.



APPENDIX III

SOME DIGITAL SIGNAL PROCESSING ACTIVITIES AT M. I. T.

A. V. OPPENHEIM *

Department of Electrical Engineering and Research Laboratory of Electronics - M. I. T. Cambridge, Mass. U.S.A.

To be presented at the Conference on SIGNAL PROCESSING University of ERLANGEN, Germany - April 4 - 6 1973

1. INTRODUCTION

Over the past five to ten years, digital signal processing has been a growing field in terms of the development of new techniques and hardware and their application to a broad class of problems [1]. Speech processing has traditionally been an important area of application for digital signal processing and the growing importance of digital signal processing techniques to speech processing is evident [2]. Similarly these techniques are increasingly important in seismic data analysis, mechanical vibration analysis, biomedical signal processing, radar and sonar systems and picture processing. This growth has resulted partly from a succession of theoretical developments such as the Fast Fourier Transform (FFT) algorithm [3], cepstral analysis and homomorphic filtering [4,5], digital filter design, signal analysis by linear prediction [6] and a long list of others. In addition the developments in integrated circuit technology have led to the implementation of increasingly sophisticated special purpose digital signal processing hardware [7] and general purpose digital signal processing computers [8].

2. DIGITAL SIGNAL PROCESSING ACTIVITIES AT M.I.T.

Digital signal processing is an active research area at M.I.T. There is a digital signal processing research group in the Research Laboratory of Electronics and another at Lincoln Laboratory. In addition there are a number of research groups which are not directly involved in research on digital signal processing but rely heavily on digital signal processing techniques to pursue their main research interests. There is generally considerable interaction among all of these groups. In this talk I will not try to be comprehensive in describing the research projects in progress, and for the most part will confine my comments to a few research areas presently being investigated by my group in the Research Laboratory of Electronics. Hopefully these comments will suggest the flavor of some of our digital signal processing activities at M.I.T..

* This work was supported in part by the National Science Foundation under grant Gk-31353 and in part by the Advanced Research Projects Agency monitored by ONR under contract N 00014-67-A-0204-0064. Presently a Guggenheim Fellow at Ecole Nationale Supérieure d'Electronique et de Radioélectricité - Grenoble, France.

The general scope of our activities includes the investigation of a number of theoretical issues and new digital signal processing techniques and the application of these techniques to speech and picture processing. As an example, over the past several years we have been interested in models for representing arithmetic errors due to finite word length arithmetic [9]. More recently we have been interested in the analysis and development of digital filter structures with low sensitivity to finite register length effects [10]. Our approach is to try to develop and utilize a network theory to obtain new filter structures. Following the approach taken by Fettweis [11] we have developed a number of network properties of digital filter including some large-scale sensitivity formulas [12]. In addition we are investigating a number of computational methods for sensitivity analysis of digital filters [13].

Another project has been directed toward the discrete representation of continuous-time signals [14]. This work has led among other things to a notion referred to as digital frequency warping which has potential application to unequal resolution and vernier spectrum analysis as discussed in detail in another paper at this conference [15]. We are presently exploring also the application of this technique to some speech processing problems.

In speech processing our interest is primarily in the application of digital signal processing techniques for bandwidth compression and coding and for parameter extraction for speech recognition, speaker verification etc. Much of our approach to speech processing has been based on homomorphic analysis of speech [5]. Recently we have been investigating and utilizing the techniques of linear prediction for speech analysis [6,16,17].

A different class of problems which are becoming increasingly important and interesting are those involving multi-dimensional digital signal processing. In image processing for example, digital filtering is important in such problems as reducing noise, inverse filtering, matched filtering for pattern recognition, contrast enhancement and dynamic range compression, etc. [18]. In contrast with one-dimensional filters however, the problem of design of multi-dimensional digital filters is for the most part unresolved. The difficulties in multi-dimensional filter design arise from the fact that the system function is a ratio of multi-dimensional polynomials. Factorization theorems for multi-dimensional polynomials do not exist as they do for one dimensional polynomials and consequently multi-dimensional system functions cannot be simply characterized in terms of poles and zeros. Thus stability is considerably harder to guarantee. Furthermore in one dimension, the approximation problem is generally based on the design of a squared magnitude function which is then factored. In multi-dimensions, it is not generally known how to specify a squared magnitude function. Furthermore, the general theory of approximation by multi-dimensional polynomials is considerably less developed. Thus the general area of multi-dimensional digital filter design presents a number of interesting research problems which we are presently pursuing.

Another project involving multi-dimensional signal processing is concerned with the digital reconstruction of multi-dimensional signals from their projections [19,20]. In many applications a set of projections of an N-dimensional object into (N-1) dimensions are available from which it is useful to reconstruct the original object. X-ray photographs for example represent two dimensional projections of the three dimensional object which has been X-rayed.

The basis for most of the algorithms which perform reconstructions is a theorem which is referred to as the projection-slice theorem. The theorem states that the projection at any orientation of an N-dimensional object has a Fourier transform which is an (N-1) dimensional slice at the same orientation of the N-dimensional Fourier transform of the object. On the basis of this each projection provides a slice of the Fourier transform and if enough slices are obtained the Fourier transform and hence the original object can be reconstructed.

Based on formulating the problem in digital signal processing terms, a number of new algorithms have been developed which appear, based on several reconstruction examples, to be superior to previous algorithms. Furthermore a number of new theoretical results have been developed relating to the minimum number of projections necessary for exact reconstruction. It has been shown in particular that taken at the right orientation exact reconstruction can theoretically be carried out from only a single projection. While reconstruction based on this theorem is too sensitive to apply in practice the theorem appears to have potential application to multi-dimensional filter design. Some of the more practical reconstruction algorithms have been applied to the reconstruction of a section of a leg bone from real X-ray data.

The above comments complete the brief survey of some of our digital signal processing activities. I would like to conclude with a short description of the computer facilities which we use.

3. FACILITIES

There are two primary facilities that we have available to us : a Digital Equipment Corporation PDP.9 computer at the Research Laboratory of Electronics and a high speed signal processing computer, the Fast Digital Processor (FDP) at Lincoln Laboratory [8]. The PDP.9 computer is a standard medium size computer with 18 bit fixed point arithmetic. In our facility we have 24K of core memory, two 250K word discs, four dec tapes a graphics display, X-Y tablet, display processor and point display, and high speed printer. In addition there is peripheral equipment interfaced for audio input and output. In general the facility is used in an interactive way.

The FDP is a special purpose programmable signal processor designed and built at Lincoln Laboratory and is interfaced to a Univac 1219 computer. It utilizes a high speed memory and operates with a memory cycle time of 250 nanoseconds. It has a highly parallel configuration with four arithmetic elements which communicate with each other and operate in parallel. In general it can implement signal processing algorithms on the order of 50-100 times faster than the PDP.9. For example, the implementation of an FFT with 256 complex or 512 real points requires on the order of 2 msec. This corresponds to real time spectral analysis at a 100 - 200 KHZ input sampling rate. A two-dimensional Fourier transform of a 256 by 256 point picture requires approximately 1 second. In general the speed of the FDP permits real time digital signal processing for a wide variety of problems. For many others such as picture processing it offers the possibility of highly interactive processing.

[1] B. GOLD, C. RADER
Digital processing of signals
Mc Graw Hill, New-York (1969)

[2] J. FLANAGAN
Speech analysis, synthesis and perception (2nd ed.)
Springer-Verlag, New-York (1972)

[3] J.W.COOLEY, J.W. TUKEY
An algorithm for the machine calculation of complex Fourier series
Mathematics of computation, 19 n° 90 pp. 297-301 (1965)

[4] B. BOGERT, M. HEALY, J. TUKEY
The quefrency alanysis of time series for echoes
Proc. Symp. Time Series Analusis - M. ROSENBLATT ed., pp. 209-243
Wiley, New-York (1963)

[5] A. OPPENHEIM, R. SCHAFER, T. STOCKHAM
Non linear filtering of multiplied and convolved signals
Proc. I.E.E.E., 56 pp. 1264-1291 (aug. 1968)

[6] B.S. ATAL, S. HANAUER
Low bit-rate speech transmission by linear prediction of speech signals
J. Acoust. Soc. Am., 50 pp. 637-655 (1971)

[7] L. JACKSON, J. KAISER, H. McDONALD
An approach to the implementation of digital filters
I.E.E.E. Trans. on Audio and Electroacoustics, AU.16 pp. 413-421 (sept.1968)

[8] B. GOLD, I. LEBOW, C. RADER, P. McHUGH
The FDP ; a fast programmable signal processor
I.E.E.E. Trans. on Computers, 20 pp. 33-38 (1971)

[9] A. OPPENHEIM, C. WEINSTEIN
Effects of finite register length in digital filtering and the Fast Fourier Transform - Proc. I.E.E.E., 60 n° 8 pp. 957-976 (aug. 1972)

[10] R. CROCHIERE
Digital Ladder structures and coefficient sensitivity
I.E.E.E. Trans. on Audio and Electroacoustics, AU.20 (oct. 1972)

[11] A. FETTWEIS
A general theorem for signal-flow networks, with applications
Arch. Elek. Ubertragung, 26 pp. 557-561 (1971)

[12] R. CROCHIERE
Some network properties of digital filters
Proc. of 1973 I.E.E.E. Intern. Symp. on Circuit Theory
Toronto, Canada (apr. 1973)

[13] R. CROCHIERE
Computational methods for sensitivity analysis of digital filters
Quaterly Progress Report n° 109, Res.Lab.of Electronics,M.I.T. (apr. 15, 1973)

[14] A. OPPENHEIM, D. JOHNSON
Discrete representation of signals
Proc. of the I.E.E.E., 60 n° 6 pp. 681-691 (june 1972)

[15] C. BRACCINI, A. OPPENHEIM
Applications of digital frequency warping to unequal bandwidth and Vernier spectrum analysis
Proc. Conf. on Signal Processing - Univ.of Erlangen, Germany (apr. 4-6, 1973)

[16] V. ZUE
Speech analysis by linear prediction - Quaterly Progress Report n° 105
Research Lab. of Electronics, M.I.T. pp. 133-142 (apr. 15, 1970)

[17] M. PORTNOFF, V. ZUE, A. OPPENHEIM
Some considerations in the use of linear prediction for speech analysis
Quater.Prog.Rep.n° 106-Res.Lab.of Electronics, M.I.T. pp.141-149 (jul.15,1972)

[18] T. HUANG, W. SCHREIBER, O. TRETIAK
Image processing
Proc. I.E.E.E. 59 n° 11 pp. 1586-1609 (nov. 1971)

[19] R. MERSEREAU
Recent advances in the theory of reconstructing multidimensional signals from projections
Quarterly Progress Report n° 105, Research Laboratory of Electronics, M.I.T. pp. 169-183 (apr. 15, 1970)
[20] R. MERSEREAU
Digital reconstruction of multidimensional signals from their projections
Sc. D. Thesis M.I.T., Department of Electrical Engineering (feb. 1973).

* * *